Task description

Run few replications of the experiment by using different seeds (integer numbers). You can use the pre-prepared acrobot.ini file included in the ./xacrobot directory. While the program is running check the source code of the environment available from the https://gym.openai.com/envs/ website to figure out the content of the observation vectors, the content of the action vector, and the way in which the reward is calculated. Plot performance across generations and then observe the behavior of evolved robots.

Solution

Issue

An error occurred in program file "policy.py", especially in the definition of Class Object GymPolicyDiscr at the 348 row which is showed in figure 1. Instead of $rew+ = rew$, the following equation should be rewritten

$$rew+ = r \tag{1}$$

```
while t < self.maxsteps:
    # Copy the input in the network
    self.nn.copyInput(np.float32(self.ob))
    # Activate network
    self.nn.updateNet()
    # Convert the action array into an integer
    action = np.argmax(self.ac)
    # Perform a step
    self.ob, r, done, _ = self.env.step(action)
    # Append the reward
    rew += rew
    t += 1
    if render:
        if (self.test == 1):
            self.env.render()
            time.sleep(0.05)
```

Figure 1: Issue in file \opt\evorobotpy\bin\policy.py

Results

The initial parameters for systems are:

- Seed 3: the default "acrobot.ini";
- Seed 5: the default "acrobot.ini";

- Seed 7: "acrobot.ini" with maxmsteps = 5, architecture = 0.

The results for different seeds:

- Seed 3: gen 822, eval 20011663, bestfit -84.00, bestgfit -86.40, centroid -87.20, bestsam -86.80, avg -93.23, weightsize 0.35, runtime 8422.32;
- Seed 5: gen 562, eval 20030245, bestfit -116.60, bestgfit -118.80, centroid -134.40, bestsam -118.60, avg -148.95, weightsize 0.37, runtime 8573.65;
- Seed 7: gen 120, eval 5040000, bestfit -200.00, bestgfit -200.00, centroid -200.00, bestsam -200.00, avg -200.00, weightsize 0.31, runtime 2331.02;

The result figure 2 is the variation of performance across generations.

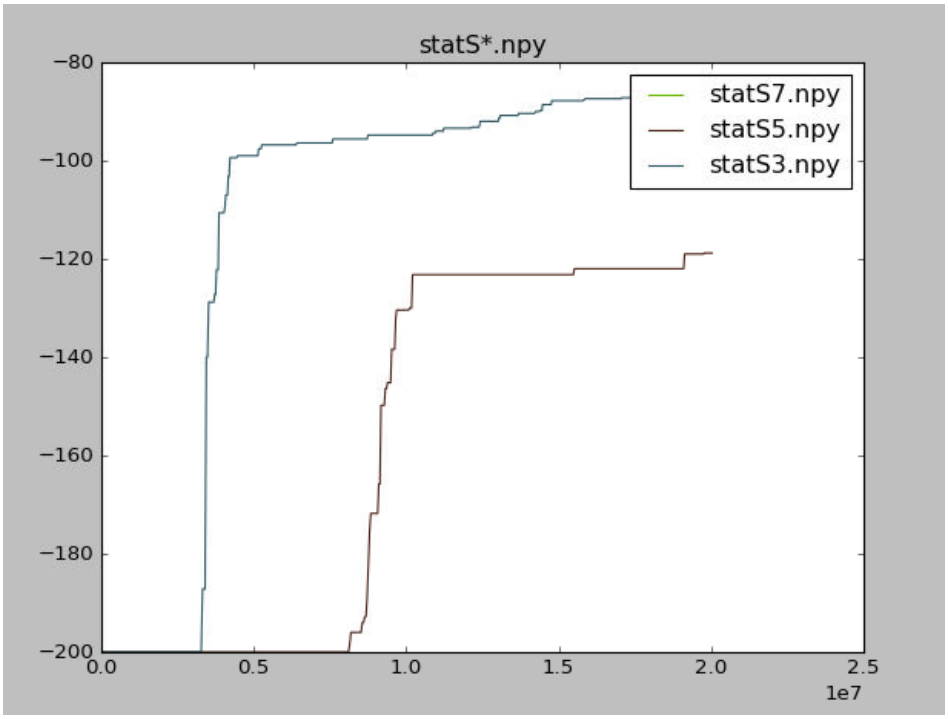

Figure 2: The graph of the variation of performance across generations

Average Generalization: -135.07 +-47.78 (3 S*.fit files).